DOCUMENT RESUME

ED 199 290 TM 810 201

AUTHOR Johnstone, Whitcomb G.
TITLE A Technique for Relating School Variables to Individual Student Achievement.
PUB DATE Apr 81
NOTE 13p.: Paper presented at the Annual Meeting of the American Educational Research Association (65th, Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement: Average Daily Attendance: Elementary Education: Grade 5: *Mathematical Formulas: *Predictor Variables: *Regression (Statistics): Resource Teachers: Socioeconomic Status
IDENTIFIERS *Criterion Variables: *Situational Regression Technique

ABSTRACT

When the school or district is the unit of analysis in school effects research, the outcome of interest is mean achievement, and the variance explained by school or district level predictors is between schools or between districts. When the individual is the unit of analysis, the outcome of interest is individual achievement and the variance explained by school or district level predictors is between individuals. A technique called situational regression (SR) is described which provides a link between these levels of analysis. SR involves regressing a criterion variable measure at the individual level on one or more predictor variables measured at the group level. Predictors are measured characteristics of the groups e.g., schools, to which individuals belong and may include socioeconomic status (SES), percent daily attendance at the school (ATTEN) and the number of resource teachers assigned to the school (REST). Data illustrating the SR technique were taken from composite student achievement scores of all grade 5 students in one district and regressed on SES, ATTEN, and REST. These three school level variables together accounted for almost 17 percent of the individual variance in achievement scores and represented nearly 70 percent of the between school variation in achievement. (RL)

A TECHNIQUE FOR RELATING SCHOOL VARIABLES TO INDIVIDUAL STUDENT ACHIEVEMENT

Whitcomb G. Johnstone
Fort Worth Independent School District

Paper presented at the annual meeting of the American Educational Research Association in Los Angeles, April, 1981

INTRODUCTION

The unit of analysis in school effects research is sometimes the individual and sometimes the school or school district (Spady, 1973; Bryant, Glazer, Hansen, and Kirsch, 1974; Centra and Potter, 1980). When the school or district is the unit of analysis the outcome of interest is mean achievement and the variance explained by school or district level predictors is between schools or between districts. When the individual is the unit of analysis the outcome of interest is individual achievement and the variance explained by school or district level predictors is between individuals. Translation between these levels has been difficult. The technique described in this paper provides a link between these levels of analysis.

The technique, which I will call "situational regression" for brevity, was developed by Demaree, James, and Hater (1979; also James, Demaree, and Hater, 1980) in the context of organizational research where questions often arise about the relationship of some characteristic of work-groups to individual productivity, absenteeism or some other criterion. In general, situational regression involves regressing a criterion variable measure at the individual level on one or more predictor variables measured at the group level. The predictors are usually measured characteristics of the groups to which individuals belong, not within

group aggregates of individual characteristics, though these are not precluded.

SITUATIONAL REGRESSION

Suppose we have access to individual student achievement data for all elementary schools in a particular district and are interested in the relationship between the composite achievement scores of fifth grade students, Y, and the SES of the school attendance areas, X. The SES measure is not aggregated over students, but represents the average SES of the census tracts overlapping each attendance area. In this case SES is considered a measured school level variable that displays no natural within school variation. Many of the predictors used in school effects research are similar in this regard, e.g. school size, per pupil expenditure, staff size, etc.

Rather than aggregate the achievement data to the level of the predictor, in situational regression each student is assigned the SES score characteristic of his or her school. The regression is then performed in the individual sample. Since this example involves only one predictor and one criterion, the proportion of achievement variance accounted for by school SES is r^2_{YX} over all students in the individual level sample. This r^2_{YX} represents the degree of linear relationship between individual achievement and school SES.

Demaree, et al., have shown that:

$$r^2_{YX} = r^2_{\bar{Y}X} \frac{\sigma^2_{\bar{Y}}}{\sigma^2_Y}$$

$$= r^2_{\bar{Y}X} \ \text{eta}^2_y. \quad (1)$$

When X is a school level predictor the variation it accounts for in Y in the individual sample is some portion of the total variation in Y accounted for by group membership. This portion is equal to the variance in the weighted group Y means accounted for by the school level predictor.

From equation (1) it is clear that a group level variable cannot account for a greater proportion of the individual level criterion variance than is associated with differences between groups. The maximum proportion of variance that can be accounted for by r^2_{YX} is eta^2_Y and will occur when $r^2_{\bar{Y}X} = 1$. This means that $r^2_{\bar{Y}X}$ is the proportion of total between group variation (eta^2_Y) accounted for by r^2_{YX}. This can be seen by reexpressing equation 2:

$$r^2_{\bar{Y}X} = \frac{r^2_{YX}}{eta^2_Y} . \quad (2)$$

These relationships may be extended to cases involving more than one school level predictor variable, i.e.:

$$R^2_{YX} = R^2_{\bar{Y}X} \; eta^2_Y, \quad (3)$$

where X represents a set of p school level predictor variables. To determine the extent to which the variation in Y accounted for by X is associated with between group differences we reexpress the equation as:

$$R^2_{\bar{Y}X} = \frac{R^2_{YX}}{eta^2_Y} . \quad (4)$$

Given that the relationships between the $\bar{Y}$ and the X are linear, $1-R^2_{\bar{Y}X}$ represents the proportion of between group variance in Y not accounted for by X. For example, Bryant, Glazer, Hansen and Kirsch (1974) present data taken from the Coleman report indicating that 10.32 percent of the variation in verbal achievement among northern white sixth grade students was accounted for by schools. The total achievement variance accounte. for by school level indices of student body characteristics, facilities and curriculum, and teacher body characteristics was 7.35 percent. What proportion of the variation between schools was accounted for by these three sets of variables? From equation 3 we have

$$R^2_{\bar{Y}X} = \frac{R^2_{YX}}{eta^2_Y} = \frac{.0735}{.1032} = .7122.$$

Approximately 71 percent of the between school variance was accounted for by the variables comprising the three sets of school level indices. About 29 percent of the between school achievement variation for white northern sixth grade students was not accounted for by the relationship of individual achievement to the three school level indices.

The link that situational regression provides between the individual level and the group level of analysis is the relationship between R^2_{YX} and $R^2_{\bar{Y}X}$. Many authors have noted that regression analyses using group means on Y amount to between group analyses (e.g. Bryant, et al., 1974; Brookover, Beady, Flood, Schweitzer, and Wisenbaker, 1979). Demaree, et al., have demonstrated that in research with intact groups the regression of individual outcomes on group characteristics

is also a between group analysis. The maximum variation any group level variable can account for in an individual level criterion variable is limited by the total criterion variance accounted for by group membership, and by the degree of linear relationship of the group criterion means to the group level predictors.

LIMITATIONS

In part, the usefulness of the situational regression technique hinges on the usefulness of eta_Y^2 as an indicator of the degree of between group variation. Glass and Hakstian (1969) correctly pointed out that the interpretation of measures like eta_Y^2 in fixed effects ANOVA designs is problematic. The problems arise from the arbitrary representation of treatment levels in such designs. The use of eta_Y^2 in school effects research employing situational regression is subject to such criticism to the extent that schools (representing levels in a one-way design) are arbitrarily selected. If, however, schools are selected at random or if, as is the case in the example to follow, all schools of interest are represented in the analysis, the use of eta_Y^2 as a measure of the total between school criterion variation is appropriate.

When group Ns, the n_i, differ widely the larger groups will be weighted more in determining R_{YX}^2. This may cause problems if the n_i are not representative of group membership in the overall population to which generalization is intended.

It is important that the number of group level predictors be smaller than the number of groups. In the limiting case, if there are k groups and $p = k - 1$ group level

predictors, the set of predictors will serve only to identify group membership. In such cases the group level variables provide little information about the sources of between group differences. Demaree, et al. suggest the rule of thumb that $R^2_{YX} > .50$ and $k > .5\ p$ hold before attaching conceptual significance to R^2_{YX}.

REAL DATA EXAMPLE

The data for this illustration of the situational regression technique were taken from the achievement test records of fifth grade students at all elementary schools in a single district. The records were arranged by school, but randomly within schools. A sample was drawn by selecting every other record. The total N was 2277 students, the number of schools was 54, and the within school n_i ranged from 20 to 100. The sampling procedure insured that the n_i would approximate their proportions in the population of the district's fifth graders.

The composite student achievement scores were regressed in stepwise fashion on the following school level variables after assigning each student his or her school's score on the variable:

1. SES, the average SES of the 1970 census tracts overlapping the school's attendance area.
2. TRANS, the rate of transiency at the school defined as the number of entrants + number of leavers divided by total enrollment.
3. ATTEN, the percent attendance at the school defined as the ratio of average daily attendance to average daily membership, taken as a percent.

4. PTR5, the pupil-teacher ratio in grade five.

5. REST, the number of resource teachers assigned to the school.

The results of the stepwise procedure are given in Table 1. The three variables selected for inclusion in the equation on the basis of partial F tests were SES, REST, and ATTEN. It should be noted that the partial F test for increments to R^2 are the same whether the predictor variables are group level or individual level.

The interpretation of R^2_{YX}, $R^2_{\bar{Y}\bar{X}}$, and $1 - R^2_{\bar{Y}\bar{X}}$ is relatively straightforward. First of all we note that $eta^2_Y = .237$, that is about 24 percent of the variance in achievement is related to differences between schools. Of the school level variables included in the analysis the socio-econimic status of the attendance area (SES), the number of resource teachers assigned the school (REST), and the percent attendance at the school (ATTEN) together accounted for almost 17 percent of the individual variance in achievement scores. This represents nearly 70 percent of the between school variation in achievement, leaving 30 percent unaccounted. According to Demaree, et al. this might indicate a need to include other school level varibles or to include school means on individual difference variables.

DISCUSSION

There have been many criticisms of school effects research, some of which were summarized by Centra and Potter (1980):

> Between-school studies have not been successful in identifying school (or district) characteristics that are highly related to how

> much students learn. Although schools apparently do make some difference, most of the variability in student achievement is related to student social class or to within-school factors. (p. 281)

The authors point out that within-school factors offer a more promising line of research, listing such things as teacher characteristics, administrative organization of the school, instructional grouping, etc. Whether the school district, school building, or classroom is the level of analysis, however, the degree of relationship of any characteristic of the group to an individual level criterion is limited by the degree to which variation in the criterion is related to group membership.

The situational regression technique provides a means to move between levels of analysis. For example, a school administrator told that mean achievement in the district's schools correlates .75 with building level ADA might start an expensive TV ad campaign to promote attendance. If, however, this person knows that the total variation in achievement accounted for by schools is .10, he or she might reconsider spending a lot of money on something that accounts for only five percent of the achievement variance between students.

In summary, situational regression provides a means for assessing the relationship between an individual level criterion variable and one or more group level predictor variables. The individual criterion variance accounted for by the group level predictors was shown to be the product of the weighted group mean criterion variance accounted for by the predictors, times the total criterion variance accounted

for between groups. The maximum amount of criterion variance in the individual sample that can be accounted for by a group level predictor is limited by the degree of between group variation in the criterion. Situational regression is most useful in cases where the predictors are measured characteristics of groups, having no individual level counterpart. In such cases sources of between group variation in the criterion variable may be explored in terms of the group level predictors that account for most variation in the sample of individuals.

REFERENCES

Brookover, W., Brady, C., Flood, P., Schweitzer, J., and Wisenbaker, J. School Social Systems and Student Achievement: Schools Can Make a Difference. New York: Praeger Publishers, 1979.

Bryant, E., Glazer, E., Hansen, M., and Kirsch, A. Associations Between Educational Outcomes and Background Variables: A Review of Selected Literature. Denver, Co.: NAEP, 1974.

Centra, J. and Potter, D. School and teacher effects: An interrelational model. Review of Educational Research, 1980, 50(27), 273-291.

Demaree, R., James, L., and Hater, J. A statistical rationale for relating situational attributes and individual differences (IBR Report 79-1). Fort Worth, Texas: Institute of Behavioral Research, Texas Christian University, 1979.

Glass, G. and Hakstian, A. R. Measures of Association in comparative experiments: Their development and interpretation. American Educational Research Journal, 1969, 6 (3), 403-414.

James, L., Demaree, R., and Hater, J. A statistical rationale for relating situational variables and individual differences. Organizational Behavior and Human Performance, 1980, 25, 354-364.

Spady, W. The impact of school resources on students. In Kerlinger, F. (ed), Review of Research in Education, Vol. 1. Itasca, Ill.: F. E. Peacock Publishers, Inc., 1973.

TABLE 1

STEP	VARIABLE ENTERED	CUMULATIVE R^2_{YX}	R^2 CHANGE	F	CUMULATIVE R^2_{YX}	CUMULATIVE $1-R^2_{YX}$
1	SES	.141	------	375.25*	.599	.401
2	REST	.156	.014	38.58*	.659	.341
3	ATTEN	.165	.009	24.14*	.697	.303

$$eta^2_y = .237$$